

CONCORSO N. 286/2023 - PROVA SCRITTA

TRACCIA 1

INDICAZIONI PER I/LE CANDIDATI/E

La prova consiste in **4 quesiti** a risposta aperta. I/le candidati/e rispondono ai quesiti seguendo le indicazioni e senza apporre nessun segno di riconoscimento.

Il tempo a disposizione per lo svolgimento della prova è pari a **90 minuti** dal segnale di VIA. È possibile consegnare prima del termine previsto.

Chi intende ritirarsi a prova già iniziata, può lasciare l'aula dopo aver restituito alla Commissione tutto il materiale ricevuto e firmato l'apposito modulo di ritiro dalla prova.

Domanda 1 [massimo 8 punti]

Un'Università dispone al suo interno di un ufficio Sviluppo Software, che ha sviluppato un applicativo per la gestione delle presenze del personale dipendente. Questo applicativo archivia i propri dati su di un DBMS locale, basato su Microsoft SQL Server. Il management dell'Università deve prendere la decisione se spostare il servizio DBMS verso un servizio database, erogato in modalità PaaS, oppure se mantenerlo "on premises". Al fine di aiutare il processo decisionale, si elenchino vantaggi e svantaggi di ciascuna opzione.

Domanda 2 [massimo 8 punti]

I ricercatori di uno dei dipartimenti dell'Università hanno la necessità di utilizzare un gran numero di pacchetti e librerie software di analisi dati scientifici, sostanzialmente diversi tra utente e utente; talvolta si tratta di configurazioni incompatibili tra loro, che vanno perciò "isolate" e rese indipendenti, pur coesistendo sulla stessa piattaforma. Si descriva e giustifichi un sistema di gestione dei pacchetti e librerie software adatto a questo caso d'uso.

Domanda 3 [massimo 8 punti]

Su di una NAS abbiamo a disposizione un sistema di "backup" basato su snapshot a livello di blocchi, che è possibile schedulare. Come può aiutarci a proteggere i dati sia dalla cancellazione accidentale da parte dell'utente, sia da attacchi (es.: ransomware)? È possibile immaginare di aggiungere componenti all'architettura se ritenuto necessario.

Domanda 4 [massimo 6 punti]

Descriva il candidato l'architettura di un cluster di calcolo ad alte performance (HPC), con particolare riguardo al gestore delle code (scheduler), ai ruoli dei membri del cluster (worker node, head node, login node).

CONCORSO N. 286/2023 - PROVA SCRITTA

TRACCIA 2

INDICAZIONI PER I/LE CANDIDATI/E

La prova consiste in **4 quesiti** a risposta aperta. I/le candidati/e rispondono ai quesiti seguendo le indicazioni e senza apporre nessun segno di riconoscimento.

Il tempo a disposizione per lo svolgimento della prova è pari a **90 minuti** dal segnale di VIA. È possibile consegnare prima del termine previsto.

Chi intende ritirarsi a prova già iniziata, può lasciare l'aula dopo aver restituito alla Commissione tutto il materiale ricevuto e firmato l'apposito modulo di ritiro dalla prova.

Domanda 1 [massimo 8 punti]

Un'Università dispone al suo interno di un ufficio Sviluppo Software, che ha sviluppato un applicativo per la gestione delle presenze del personale dipendente. Questo applicativo archivia i propri dati su di un DBMS locale, basato su Microsoft SQL Server. Il management dell'Università prende la decisione di spostare l'applicativo, e con esso il database, verso un cloud provider di tipo "public" (come Azure, AWS, GCE, ...). Si ipotizzi di dover delineare un piano di migrazione dei contenuti della base dati da un'istanza "on premises" di SQL Server ad un'istanza "on cloud" di SQL Server, erogata in modalità PaaS. È possibile impiegare tutti gli step intermedi ritenuti necessari.

Domanda 2 [massimo 8 punti]

Si spieghi la differenza tra virtualization e containerization. Si illustrino inoltre i criteri che si possono utilizzare per scegliere se mettere a disposizione all'utenza un applicativo scientifico di analisi dati tramite il primo o il secondo approccio.

Domanda 3 [massimo 8 punti]

Un moderno sistema NAS viene valutato, oltre che in base alle caratteristiche più propriamente "funzionali", come capacità massima installabile, throughput e tempo di risposta, anche in base a requisiti "non funzionali", come affidabilità, scalabilità, disponibilità e consistenza. Tramite quali tecniche è possibile migliorare queste caratteristiche in uno storage di rete?

Domanda 4 [massimo 6 punti]

Un cluster di calcolo HPC è composto da un certo numero di computer (nodi), dei quali alcuni con compito di coordinamento (head node), dove viene eseguito lo scheduler, altri che eseguono i lavori sottomessi al cluster (worker node), e infine alcuni che danno la possibilità all'utente del cluster di collegarsi ed accodare lavori alle code di esecuzione (login node). Spieghi il candidato quali sono i vantaggi della distinzione tra ruoli diversi per differenti nodi di calcolo all'interno dell'architettura del cluster, rispetto ad una soluzione "ibrida" in cui ciascun nodo assume su di sé tutti e tre i compiti indicati in precedenza.

CONCORSO N. 286/2023 - PROVA SCRITTA

TRACCIA 3

INDICAZIONI PER I/LE CANDIDATI/E

La prova consiste in **4 quesiti** a risposta aperta. I/le candidati/e rispondono ai quesiti seguendo le indicazioni e senza apporre nessun segno di riconoscimento.

Il tempo a disposizione per lo svolgimento della prova è pari a **90 minuti** dal segnale di VIA. È possibile consegnare prima del termine previsto.

Chi intende ritirarsi a prova già iniziata, può lasciare l'aula dopo aver restituito alla Commissione tutto il materiale ricevuto e firmato l'apposito modulo di ritiro dalla prova.

Domanda 1 [massimo 8 punti]

Uno dei vantaggi dei sistemi "cloud-based" rispetto a sistemi "on-premises" è la semplicità con cui è possibile aumentare le caratteristiche "non funzionali" (affidabilità, disponibilità, scalabilità, eccetera) della propria piattaforma, rispetto appunto ad un sistema tradizionale, basato su di un datacenter "locale": è infatti sufficiente modificare i parametri del servizio per ottenere l'effetto desiderato. Si chiede di spiegare:

- quali particolarità dei sistemi cloud consentono di implementare queste capacità;
- quali vantaggi ulteriori, e quali svantaggi, può offrire oltre a quelli già elencati, un approccio "multi-cloud", ossia che impiega le risorse di cloud multipli simultaneamente per migliorare le caratteristiche della propria piattaforma.

Domanda 2 [massimo 8 punti]

Un'Università è in procinto di migrare parte della sua infrastruttura informatica verso un cloud di tipo "public", con servizi erogati in modalità IaaS. Si vorrebbe approfittare della migrazione per ridefinire l'architettura complessiva secondo il paradigma "infrastructure-as-a-code" (IaaC). Si illustri brevemente in cosa consiste questo paradigma e quali siano i vantaggi e gli svantaggi rispetto ad un'impostazione "tradizionale" che preveda la configurazione come risultato delle opzioni scelte da un gruppo di amministratori tramite apposite interfacce di gestione.

Domanda 3 [massimo 8 punti]

Illustrare vantaggi e svantaggi di una soluzione di archiviazione file con condivisione tramite rete (share) centralizzata tipo NAS rispetto ad un'archiviazione dei dati locali all'infrastruttura che li ha prodotti, con particolare riguardo alle procedure di gestione degli accessi e dei permessi, dei backup automatizzati e della sicurezza dei dati.

Domanda 4 [massimo 6 punti]

I più potenti supercomputer al mondo sono spesso costituiti da unità elaborative più semplici, interconnesse tra di loro; queste unità sono fortemente omogenee, e sono sistemi di elaborazione sviluppati esclusivamente per quello specifico impiego. Come si compara, in termini di vantaggi e svantaggi, questo scenario con quello di un cluster di calcolo che impiega hardware "comune" (cioè, server "general purpose", che possono quindi essere impiegati anche per altri compiti), ed eterogeneo (per esempio, perché diverse "generazioni" di dispositivi vengono aggiunte in momenti successivi al cluster)? Quali tecniche possono facilitare la gestione di quest'ultimo caso?